PATENTTI- JA REKISTERIHALLITUS
NATIONAL BOARD OF PATENTS AND REGISTRATION

PCT/FI00/00819

Helsinki 7.8.2000

REC'D 06 OCT 2000
WIPO PCT

E T U O I K E U S T O D I S T U S
P R I O R I T Y D O C U M E N T

FJU

Hakija Applicant	**Galilaeus Oy** **Piispanristi**
Patenttihakemus nro Patent application no	**19992085**
Tekemispäivä Filing date	**29.09.1999**
Kansainvälinen luokka International class	**C12N**
Keksinnön nimitys Title of invention	

"The gene cluster involved in aclacinomysin biosynthesis, and its use for genetic engineering"
(Aklasinomysiinin biosynteesiin liittyvä geeniryhmittymä ja sen käyttö-geenitekniikassa)

Täten todistetaan, että oheiset asiakirjat ovat tarkkoja jäljennöksiä patentti- ja rekisterihallitukselle alkuaan annetuista selityksestä, patenttivaatimuksista, tiivistelmästä ja piirustuksista.

This is to certify that the annexed documents are true copies of the description, claims, abstract and drawings originally filed with the Finnish Patent Office.

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

Eija Solja
Eija Solja
Apulaistarkastaja

Maksu 300,- mk
Fee 300,- FIM

Osoite: Arkadiankatu 6 A, P.O.Box 1160, FIN-00101 Helsinki, FINLAND
Puhelin: 09 6939 500
Telephone: + 358 9 6939 500
Telefax: 09 6939 5328
Telefax: + 358 9 6939 5328

The gene cluster involved in aclacinomycin biosynthesis, and its use for genetic engineering

Field of the invention

This invention relates to the gene cluster for aclacinomycin biosynthesis derived from *Streptomyces galilaeus*, and the use of the genes included therein to obtain hybrid antibiotics, or to increase yields of aclacinomycins or related antibiotics.

Background of the invention

Anthracyclines are widely used anticancer agents. Seven different anthracyclines are in worldwide clinical use: daunorubicin, doxorubicin, idarubicin, epirubicin, pirarubicin, zorubicin and aclarubicin. A representative compound is doxorubicin, being the most efficient and acting on a wide array of malignancies. A variety of toxic effects, like cumulative cardiotoxicity found with doxorubicin has sometimes led to discontinuation of the treatment. Furthermore, there are some type of malignancies which do not respond to available anthracyclines. The mechanism of action of anthracyclines, reflecting to their clinical efficiencies, is not clear, although most researchers consider inhibition of topoisomerase II as a desired effect. Generation of free radicals derived from quinonic structures is suggested to be related to side effects such as cardiotoxicity. Anthracyclines have recently been reviewed by Professor Strohl and his group (1997).

Aclacinomycin A (aclarubicin) first described by Oki *et al.* (1975) is an anthracycline antibiotic produced by *Streptomyces galilaeus* ATCC 31133 and *S. galilaeus* ATCC 31615. It is active against tumor cells and exhibits alleviated toxic properties as compared with doxorubicin. However, its activity does not reach solid tumors, limiting its use in leukemia treatment. Aclarubicin differs from the other counterparts in its structure. A trisaccharide moiety, rhodosamine–2–deoxyfucose–cinerulose A is attached at C–7 by a glycosidic bond, whereas at the corresponding position of daunomycins only one sugar residue, daunosamine, is attached.

Despite the long history of anthracyclines, three decades or so, the studies on their biosynthesis are still going on, and there is further interest to obtain novel molecules for

the development of cancer chemotherapeutics. A method currently used for finding novel molecules for drug screening is genetic engineering. Cloning the genes for anthracycline biosynthesis facilitates the production of hybrid anthracyclines, as well as their use in combinatorial biosynthesis to generate novel molecules. As regards the chemical nature of anthracyclines currently in clinical use, aclarubicin has unique features which make its biosynthetic genes interesting in creating novel products.

Regarding the genes for deoxyhexose pathway, Madduri *et al.* (1998) have reported that a gene derived from avermectin biosynthesis cluster caused the production of hybrid anthracyclines altering a sugar moiety when transferred into a *S. peucetius* strain. The product obtained was epirubicin, a commercially important anthracycline. In this case a hydroxy group in the daunosamine moiety was in the opposite stereochemistry due to the action of an avermectin biosynthesis gene.

S. galilaeus has been used as the host to prepare hybrid anthracyclines using the genes derived from rhodomycin pathway from *S. purpurascens* (Niemi *et al.,* 1994) and from nogalamycin biosynthesis cluster from *S. nogalater* (Ylihonko *et al.,* 1996a). The genes for nogalamycin pathway were used to generate the hybrid anthracycline production in *S. steffisburgensis* producing typically steffimycin (Kunnari *et al.,* 1997). Previously, biosynthesis genes for actinorhodin have been expressed in *S. galilaeus*, resulting in the formation of aloesaponarin (Strohl *et al.,* 1991). These hybrid compounds were modified in the aglycone moiety. Recently, the biosynthesis genes involved in deoxyhexose pathway of nogalamycin were used to generate hybrid compounds using the *S. galilaeus* mutants as hosts (FI pat. appln No. 982295).

As shown above, *S. galilaeus* has been used as a cloning host to generate novel molecules, whereas its use to donate the genes has not been described. The identified genes involved in aclacinomycin biosynthesis include polyketide reductase gene (Tsukamoto *et al.,* 1994), aklanonic acid methyl ester cyclase (GeneBank, ACCESSION AF043550) and genes for polyketide synthase (Hutchinson and Fujii, 1995; the sequence not available).

Summary of the invention

The present invention concerns a gene cluster, most of the genes of which are derived from deoxyhexose pathway for rhodosamine, 2–deoxyfucose and/or rhodinose. The gene cluster was cloned from *S. galilaeus* ATCC 31615 and it is involved in biosynthesis of aclacinomycins.

Detailed description of the invention

The experimental procedures of the present invention include biochemical and chemical methods conventional in the art. Detailed description of the techniques not explained here are given in the manuals by Hopwood *et al.* 'Genetic manipulation of Streptomyces: a laboratory manual'. The John Innes Foundation, Norwich (1985) and by Sambrook *et al.* (1989) 'Molecular cloning: a laboratory manual'.

The publications, patents and patent applications cited herein are given in the reference list in their entirety.

The present invention concerns particularly the discovery of the gene cluster for aclacinomycin biosynthesis. The cluster, when introduced into *S. peucetius* strains caused the production of hybrid antibiotics modified in their sugar moiety.

Several strategies may be adopted to clone genes for an antibiotic. Using *E. coli* as a host for a gene library, hybridization is the most advantageous screening strategy. The probe for hybridization may be any known fragment that shows sufficient homology to the biosynthetic cluster for aclarubicin sugars, to be able to hybridize with said cluster. A DNA fragment which is identical to the desired region is preferred. Such a fragment, called Sg–dht, was obtained by PCR amplification of *S. galilaeus* chromosomal DNA, using degenerated oligonucleotides annealing to the conserved region of 4,6–dehydratase gene. 4,6–dehydratase is the first enzyme participating to a reaction series that converts a glucose molecule bound to a nucleotide into 6–deoxy sugars generally found in antibiotics. Using this probe it was possible to clone the cluster of deoxyhexose pathway

from a restricted gene library. To simplify the cloning strategy the library was prepared in a pUC-based plasmid (e.g. pBluescript or pWHM1109) replicating in *E. coli*.

The strategy to clone the genes involved in aclacinomycin biosynthesis according to the invention was in brief: Total DNA was isolated from *S. galilaeus* (ATCC 31615) and digested with several restriction enzymes that yield fragments of 10 kb in average. Restriction fragments were analyzed by Southern hybridization using a homologous DNA fragment, Sg-dht, as a probe. *Bgl*II gave a hybridized fragment of 8.5 kb, and a double digestion with *Xho*I and *Not*I gave a hybridized fragment of 7 kb. DNA digestion using (i) *Bgl*II and (ii) *Xho*I-*Not*I was carried out and the fragments were ligated to the *E. coli-Streptomyces* shuttle vector, pWHM1109, digested with *Bam*HI and to the pBluescript digested with *Xho*I-*Not*I, respectively. The ligation mixtures were introduced into *E. coli* XL1BlueMRF' that exhibits alleviated restriction-modification systems. Colonies were plated on the agar plates in the dilution to give 200 to 600 cfu (colony forming units) per plate. Well grown colonies were transferred in nylon membranes for hybridization, which was carried out using the Sg-dht probe. Six out of the 786 *Bgl*II-digested clones gave hybridization signal and 7 out of 1523 of those clones carrying *Xho*I-*Not*I fragments. Hybridization and washes were carried out in the stringent conditions of 65°C in a low salt concentration. Several techniques for the labeling of the probe and for hybridization are possible, but the procedure according to Boehringer Mannheim's "The DIG System User's Guide for Filter Hybridization" is preferred. The colonies giving hybridization signals were cultivated for plasmid isolation. The plasmids were analyzed by Southern hybridization to confirm the reliability of the colony hybridization. Plasmids containing the desired DNA fragments (Sg4 and Sg5) were designated as pSgc4 (*Bgl*II-fragment) and pSgc5 (*Xho*I-*Not*I fragment)(see Fig. 2).

The fragments, Sg4 and Sg5, were subcloned for sequencing in *E. coli* vectors pUC19 and pBluescript. In total 30 subclones were used to obtain the nucleotide sequence of Sg4 and Sg5. The sequenced cluster revealed thirteen genes involved in biosynthesis of aclacinomycins. Comparison with the sequences found in the sequence library suggested the functions as *sga*2 for an activator, *sga*3 for a dehydratase, *sga*4 for oxidoreductase, *sga*5 for dTDP-glucose 4,6-dehydratase, *sga*6 for glycosyl transferase (GTF), *sga*7 for

a putative isomerase, *sga*8 for aklaviketone reductase, *sga*9 for a putative polyketide assembler, *sga*10 for a putative cyclase, *sga*11 for aminomethylase, *sga*12 for glucose–1–phosphate thymidylyl transferase, *sga*13 for aminotransferase. The function of *sga*1 is not suggested based on similarity searches. Based on the deduced functions, nine genes are involved in glycosylation pathway. The genes involved in the formation of aglycone are *sga*8, *sga*9, and *sga*10. The activator, Sga2, may control both the glycosylation system and the formation of aklavinone via polyketide pathway.

Sg4 derived from pSgc4 was cloned in the *Streptomyces* expression vector pIJE486 (Ylihonko *et al.,* 1996b) in *S. lividans* TK24 to give pSgs4. This vector is a high copy number plasmid that replicates in several *Streptomyces* spp. (Ward *et al.,* 1986) and it contains a constitutively expressed promoter, *erm*E (Bibb *et al.,* 1985) upstream from the multiple cloning site. The plasmid pSgs4 isolated from TK24 was introduced into the *S. galilaeus* strains that are blocked in deoxyhexose pathway of aclacinomycin biosynthesis and into the *S. peucetius* mutants producing ε–rhodomycinone based on a lesion in glycosylation genes. The ability of aclacinomycin production was restored by three *S. galilaeus* mutants, H063, H054 and H065. The mutant strain H063 accumulates aklavinone and it was completely complemented by the plasmid pSgs4. Instead, H054 and H065 producing aklavinone glycosides sharing neutral sugars, but not rhodosamine, were only partially complemented by pSgs4. Surprisingly, H063 carrying pSgs4 (H063/pSgs4) was able to produce aclacinomycins two–fold to that of the wild type *S. galilaeus*. *S. peucetius* M18 and M90 which produce ε–rhodomycinone were selected to hosts for pSgs4. L–rhamnosyl–ε–rhodomycinone (El Khamed *et al.*, 1977) was obtained when pSgs4 was expressed in the mutants M18 and M90 and, in addition, M18/pSgs4 produced L–daunosaminyl–ε–rhodomycinone (Essery and Doyle, 1980). The structures were not new ones but this demonstrates the ability of the gene cluster according to the present invention to generate hybrid products in a heterologous host. To produce hybrid compounds we prefer to use E1 medium supplemented with a suitable antibiotic, in this case, thiostrepton, to maintain the selection pressure for the plasmid containing strains. The products were extracted by organic solvents and purified by chromatography to obtain the compounds in high purity for structural elucidation.

Examples to further illustrate the invention are given hereafter.

Brief description of the drawings

FIG. 1 shows the structures of aclacinomycin, daunomycin and ε–rhodomycinone.

FIG. 2 is a diagram of the gene cluster for aclacinomycin biosynthesis.

FIG. 3 describes the proposed biosynthesis pathway for sugars found in aclacinomycins.

FIG. 4 shows the structures of the hybrid compounds produced by M18/pSgs4 (**1** and **2**) and M90/pSgs4 (**2**).

EXPERIMENTAL

Materials used

Restriction enzymes used were purchased from Promega (Madison,Wisconsin, USA), Fermentas (Lithuania) or Boehringer Mannheim (Germany), alkaline phosphatase from Boehringer Mannheim, and used according to manufacturers' instructions. Proteinase K was purchased from Promega and lysozyme from Sigma. Hybond™–N nylon membranes used in hybridization were purchased from Amersham (Buckinghamshire, England), DIG DNA Labelling Kit and DIG Luminescent Detection Kit from Boehringer Mannheim. Qiaquick Gel Extraction Kit from Qiagen (Hilden, Germany) was used for isolating DNA from agarose.

Bacterial strains and their use

Escherichia coli XL1BlueMRF' (Stratagene, La Jolla, California) was used for cloning.

Streptomyces lividans TK24 was the first cloning host for gene expression. The strain was provided by prof. Sir David Hopwood, John Innes Centre, UK.

The wild type, *Streptomyces galilaeus* ATCC 31615, produces aclacinomycins. It was used here to donate the genes of the invention.

Streptomyces galilaeus H039 (Ylihonko *et al.,* 1994) produces Akv–$(Rho)_{0-3}$. It was used as an expression host for pSgs4 being more easily transformed than the other mutants or the wild type.

Streptomyces galilaeus H054 (Ylihonko *et al.,* 1994) produces Akv–Rho–dF–$(CinA)_{0-1}$, Akv–dF–dF–$(CinA)_{0-1}$ and Akv–dF–Rho–Rho. It was used as an expression host for pSgs4.

Streptomyces galilaeus H063 produces aklavinone. It is a mutant strain derived from the wild type *S. galilaeus*. H063 was used as an expression host for pSgs4.

Streptomyces galilaeus H065 produces aklavinone with neutral glycosides. It is a mutant strain derived from the wild type *S. galilaeus*. H065 was used as an expression host for pSgs4.

Streptomyces peucetius M18 and M90 producing ε–rhodomycinone are the mutants derived from *S. peucetius* var. *caesius* (ATCC 27952). They were used as expression hosts for pSgs4.

Plasmids

E. coli cloning vectors pBluescript SK (Stratagene) and pUC19 (Pharmacia, Sweden) were used for making the subclones for sequencing and pBluescript was used also as a vector of a gene library.

pWHM1109 (provided by prof CR Hutchinson, Wisconsin, USA) is a shuttle vector replicating in *E. coli* and in streptomycetes. It was used as a vector of a gene library.

pIJ486 is a high copy plasmid vector provided by prof. Sir David Hopwood, John Innes Centre, UK (Ward *et al.,* 1986).

pIJE486 (Ylihonko *et al.*, 1996b) is an expression vector containing *erm*E (Bibb *et al.*, 1985) to promote expression of the cloned genes.

Nutrient media and solutions

For cultivation of *S. galilaeus* for total DNA isolation TSB medium was used. Lysozyme solution (0.3 M sucrose, 25 mM Tris, pH 8 and 25mM EDTA, pH 8) was used to isolate total DNA. TE buffer (10 mM Tris, pH 8.0 and 1mM EDTA) was used to dissolve DNA.

TRYPTONE-SOYA BROTH (TSB)

Per litre: Oxoid Tryptone Soya Broth powder 30 g.

ISP4

Bacto ISP-medium 4, Difco; 37 g/l.

E1 Per litre in tap water:

glucose	20 g
soluble starch	20 g
Farmamedia	5 g
Yeast extract	2.5 g
$K_2HPO_4 \circ 3H_2O$	1.3 g
$MgSO_4 \circ 7H_20$	1 g
NaCl	3 g
$CaCO_3$	3 g

pH adjusted to 7.4 before autoclaving

General methods:

NMR data was collected with a JEOL JNM-GX 400 spectrometer. ^1H and ^{13}C NMR samples were internally referenced to TMS.

The anthracycline metabolites were determined by (i) HPLC (LaChrom, Merck Hitachi, pump L-7100, detector L-7400 and integrator D-7500) using a LiChroCART RP-18 column. Acetonitrile:potassium hydrogen phosphate buffer (60 mM, pH 3.0 adjusted

with citric acid) was used as a mobile phase. Gradient system starting from 65 % to 30 % of potassium dihydrogen phosphate buffer was used to separate the compounds. The flow rate was 1 ml/min and the detection was carried out at 480 nm, and (ii) by TLC using precoated Kieselgel 60 F_{254} glass plates (Merck, Darmstadt, Germany) with an elution solution of toluene:ethyl acetate:methanol:formic acid (50:50:15:3).

ISP4 plates supplemented with thiostrepton (50 μg/ml) were used to maintain the plasmid carrying cultures.

Example 1. Cloning the gene cluster for aclacinomycin biosynthesis

1.1 Selection of clones by hybridization

For isolation of total DNA, *Streptomyces galilaeus* was grown for four days in 50 ml of TSB medium supplemented with 0.5% glycine. The cells were harvested by centrifuging for 15 min (3900 x g) in 12 ml Falcon tubes, and stored at −20°C. Cells from a 50 ml culture were used to isolate DNA. 5 ml of lysozyme solution containing 5 mg/ml of lysozyme was added on the cells of each Falcon tube, and incubated for 20 min at 37°C. 500 μl of 10% SDS containing 0.7 mg of proteinase K was added on the cells, and incubated for 80 min at 62°C, another 500 μl of 10% SDS containing 0.7 mg of proteinase K was added, and incubation was continued for 60 min. The sample was chilled on ice and 600 μl of 3M NaAc, pH 5.8 was added, and the mixture was extracted with equilibrated phenol (Sigma). The phases were separated by centrifuging (1400 x g) for 10 min. The DNA was precipitated from the water phase with an equal volume of isopropanol and collected by spooling with a glass rod and washed by dipping into 70% ethanol, air dried and dissolved in 500 μl of TE–buffer.

Southern hybridization to determine suitable restriction enzymes for preparing the restricted plasmid libraries was carried out using *Bgl*II, *Xho*I, *Not*I and their combinations. A fragment of about 9 kb hybridizing with the Sg–dht probe was preferred. For hybridization 600 ng of digested *S. galilaeus* DNA was loaded onto the agarose gel and after electrophoresis, the DNA was transferred from the gel to a nylon membrane by vacuum blotting. Hybridization was carried out according to Boehringer Mannheim's

manual 'The DIG System User's Guide for Filter Hybridization'. The probe for hybridization, Sg–dht, which was used for colony hybridization as well, was obtained by amplifying a gene fragment from the *S. galilaeus* DNA which is internal to the 4,6–dehydratase gene and corresponds to the fragment of 6345 to 6861 shown in SEQ ID NO:14. PCR was used for amplification, and the sequences for the degenerated oligonucleotide primers were 5'–CSGGSGSSGCSGGSTTCATSGG–3' (forward, SEQ. ID. NO:15) and 5'–GGGWRCTGGYRSGGSCCGTAGTTG–3' (reverse, SEQ. ID. NO:16). Suitable fragments were a 9 kb *Bgl*II fragment and a 7 kb *Xho*I–*Not*I fragment.

Ten micrograms of the chromosomal DNA was digested with *Bgl*II. The DNA fragments were separated by agarose gel electrophoresis and the band of 8 to 9 kb were cut from the 0.6% low gelling temperature SeaPlaque® agarose. The DNA band was isolated from the gel using Qiagen Gel Extraction Kit. The isolated fragment was ligated to pWHM1109 plasmid vector digested with *Bam*HI and defosforylated, in the ratio of 3 moles of the insert DNA to 1 mole of the vector DNA. The ligated DNA was introduced into *E. coli* XL1BlueMRF' by electroporation. Using the whole ligation mixture 786 colonies were obtained. The colonies were grown on agar plates for at least 12 h and transferred to nylon membranes. Hybridization of colony membranes was carried out as Southern using Sg–dht as a probe. Six clones gave signal in hybridization and the corresponding colonies were plated on agar and inoculated in 3 ml of LB medium for isolation of the plasmid DNA. Southern hybridization was used to study whether the plasmids derived from the clones carried the desired insert. Four of these plasmids contained the 4,6–dehydratase gene fragment and gave the identical restriction map thus carrying the same fragment representing both orientations. The fragment was designated as Sg4 and the plasmid containing the fragment as pSgc4.

In the same manner the plasmid library representing a 7 kb *Xho*I–*Not*I DNA fragment derived from *S. galilaeus* was constructed. pBluescript was digested with *Xho*I–*Not*I and the library containing the gene fragments of around 7 kb was constructed. In total 1523 colonies were hybridized and seven turned to be the desired clone. As described above, the clones were studied for the *Xho*I–*Not*I fragment. The insert fragment was designated as Sg5 and the plasmid as pSgc5. The strain *E. coli* XL1Blue MRF'/pSgc5 obtained was deposited according to the rules of the Budapest Treaty at Deutsche Sammlung von

Mikroorganismen und Zellkulturen GmbH (DSMZ) on August 12, 1999 with the accession number DSM 12999. The fragments Sg4 and Sg5 overlap within 836 bp corresponding bases from 6181 to 7016 in SEQ ID NO:14.

1.2. Subcloning the fragments for sequencing

To determine the nucleotide sequence of the whole cluster of the Sg4 and Sg5 suitable subclones were constructed. The convenient restriction sites were used for subcloning the 14806 bp region in the plasmids pUC19 and pBluescript. Nineteen subclones were needed to sequence Sg4, and 11 subclones for Sg5.

E. coli XL1BlueMRF' cells containing the subcloned plasmids were cultivated over-night at 37°C in 5 ml of LB–medium supplemented with 50 µg/ml of ampicillin. To isolate plasmids for sequencing reactions Wizard Plus Minipreps DNA Purification System kit of Promega or Biometra Silica Spin Disc Plasmid DNA Miniprep kit of Biomedizinische Analytik Gmbh were used according to the manufacturers' instructions.

DNA sequencing was performed using the automatic ABI DNA sequencer (Perkin–Elmer) according to the manufacturer's instructions.

1.3 Sequence analysis and the deduced functions of the genes

Sequence analyses were made using the GCG sequence analysis software package (Version 8; Genetics Computer Group, Madison, Wis., USA). The translation table was modified to accept also GTG as a start codon. Codon usage was analyzed using published data (Wright and Bibb 1992).

According to the CODONPREFERENCE program the sequenced DNA fragment revealed 11 complete open reading frames (ORFs), and two 5' ends of the other ORFs (*sga*1 and *sga*13). The functions of the genes were concluded by comparing the amino acid sequences translated from their base sequences to the known sequences in the data banks. The results are shown in Table 1 referring to the sequence data given in the application.

The suggested functions for the genes match well with a proposed biosynthetic pathway of sugars of aclacinomycins (Fig. 3). The last residue in a trisaccharide moiety of aclacinomycins is rhodinose that is enzymatically converted to cinerulose. Aclacinomycin N, a precursor of aclarubicin, contains rhodinose as the third sugar residue.

Table 1.

Gene	Position	Amino acids	Deduced function	Remarks
*sga*1	-1986 compl	>662	unknown	not complete Seq.ID.NO:1
*sga*2	2523-3341	272	activator	Seq.ID.NO:2
*sga*3	3355-4659 compl	434	dehydratase	Seq.ID.NO:3
*sga*4	4821-5810	329	oxidoreductase	Seq.ID.NO:4
*sga*5	5920-6891 compl	323	dTDP-glucose 4,6-de-hydratase	Seq.ID.NO:5
*sga*6	6879-8210 compl	443	glycosyl transferase (GTF)	Seq.ID.NO:6
*sga*7	8287-9618 compl	443	putative isomerase	Seq.ID.NO:7
*sga*8	9642-10445 compl	267	aklaviketone reductase (KRII)	Seq.ID.NO:8
*sga*9	10471-10905 compl	144	putative polyketide assembler	Seq.ID.NO:9
*sga*10	11115-11894	259	putative cyclase	Seq.ID.NO:10
*sga*11	11956-12672	238	aminomethylase	Seq.ID.NO:11
*sga*12	12685-13560 compl	291	glucose-1-phosphate thymidylyltransferase	Seq.ID.NO:12
*sga*13	13783-14805	341	aminotransferase	Seq.ID.NO:13 not complete

1.4 Expression cloning in Streptomyces strains

The 8 kb *Bam*HI-*Hind*III fragment from pSgc4 was ligated in pIJE486 to give pSgs4. Plasmid pSgs4 was introduced into *S. lividans* TK24 by protoplast transformation. The strain *S. lividans* TK24/pSgs4 obtained was deposited according to the rules of the

Budapest Treaty at Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ) on August 12, 1999 with the accession number DSM 12998. The plasmid pSgs4 was isolated from the strain, and further transferred into *S. galilaeus* mutant H039. The plasmid preparate isolated from H039 was subsequently introduced into H063, H054, and H065 mutants deficient of glycosylation system of aclacinomycins. The usage of H039 as a primary *S. galilaeus* host was due to the better efficiency for the intake of foreign DNA.

S. galilaeus mutants were studied for complementation by cultivating the clones containing pSgs4 in E1 medium supplemented with thiostrepton (10 μg/ml). The products from a 500 μl sample of the culture broth were extracted with toluene:methanol (1:1) at pH 7. The metabolites from both the transformed clones and the mutants were analyzed by TLC and HPLC to find the differencies caused by pSgs4. H063 producing endogenously aklavinone was restored to aclacinomycin producer with pSgs4. No aklavinone was found in the culture broth of H063/pSgs4. However, complementation was not completed when pSgs4 was expressed in H054 and H065. Both of the mutants produce aklavinone with neutral glycosides. Incomplete complementation was presumably due to the loss of the plasmids of some bacterial cells during cultivation, or a low expression of the genes needed as an activator is not present in pSgs4.

In the same manner, pSgs4 isolated from TK24 was introduced into the *S. peucetius* mutants M18 and M90. The characteristic product for these mutants is ε–rhodomycinone. The strains M18/pSgs4 and M90/pSgs4 containing the plasmid were cultivated in E1 medium supplemented with thiostrepton (10 μg/ml), and the metabolites therein were analyzed by TLC and HPLC. Both of the clones revealed an altered production profile as compared with the products obtained from the mutants. M90/pSgs4 accumulated a glycosylated product, yielding ε–rhodomycinone as the aglycone. The compound was identified as L–rhamnosyl–ε–rhodomycinone which has been previously synthesized (CAS=63252–11–9) by El Khamed *et al.* (1977).

M18/pSgs4 produced two compounds differing from the parental strain. According to the HPLC and TLC data one compound was the same as was produced by M90/pSgs4,

L-rhamnosyl-ε-rhodomycinone, and the other one was L-daunosaminyl-ε-rhodomycinone, which was previously characterized by Essery and Doyle (1980).

Table 2: TLC and HPLC data of the hybrid products

Product	Rf-value	Retention time
ε-rhodomycinone	0.67	6.70
L-rhamnosyl-ε-rhodomycinone	0.38	5.00
L-daunosaminyl-ε-rhodomycinone	0.04	4.06

1.5 Applicability of pSgs4 for strain improvement

Since H063 was completely complemented by pSgs4, the production level of aminoglycosides was studied. For this purpose, H063/pSgs4, H063 and the wild type *S. galilaeus* were cultivated in E1 medium in the Erlenmeyer bottles for four days. Two samples of 2 ml from each culture were extracted first with toluene:methanol (1:1) in acidic conditions to remove the neutral glycosides and the aglycones. The extraction procedure was repeated until neutral glycosides and the aglycones had disappeared from the water phase. The amount of anthracycline metabolites in toluene phase was determined and is shown in Table 3. Aclacinomycins containing rhodosamine were extracted from the water phase by chloroform. Both toluene and chloroform extracts were analyzed by TLC and toluene phases contained mostly aklavinone and the degradative products. Chloroform phases contained mainly aminoglycosides, although minor amounts of the aglycones were also detected. Extracts were evaporated to dryness and subsequently dissolved into 1 ml of methanol. The amounts of anthracycline metabolites were detected by spectrophotometer at 430 nm. The amounts related to absorbance were calculated using an extinction coefficient of 13000. The results given as mg/l of cultivation broth are shown in Table 3. The production of aclacinomycins by H063/pSgs4 was at least twofold better than obtained by the wild type.

Table 3.

	Chloroform phase aminoglycoside fraction		Toluene phase aglycone fraction	
Sample	Absorbance	Concentration (mg/l)	Absorbance	Concentration (mg/l)
H063	0.401	12.6	2.956	92.3
H063/pSgs4	2.751	85.9	2.974	92.9
S. galilaeus	1.338	41.8	0.690	21.5

The ability to increase the yield of aclacinomycins by pSgs4 in the mutant H063 suggests that the genes according to the present invention are useful in strain improvement.

Example 2. Compounds generated by pSgs4

The seed culture, 180 ml of E1 culture of the plasmid containing strains, M18/pSgs4 or M90/pSgs4, was obtained by cultivating each of the strains in three 250 ml Erlenmeyer flasks containing 50 ml of E1−medium supplemented with thiostrepton (5 μg/ml) for four days at 30°C, 330 rpm. The combined culture broths (180 ml) were used to inoculate 13 l of E1−medium in a fermentor (Biostat E). Fermentation was carried out for five days at 28°C (330 rpm, aeration: 450 l/min).

The cells were harvested by centrifuging. 2.6 l of methanol was used to brake the bacterial cells. The anthracycline metabolites were extracted from methanol solution at pH 8 using 2 l of ethyl acetate and the extract was evaporated to dryness. The viscous residue was loaded onto a silica column of 4 × 10 cm and toluene:ethyl acetate:formic acid (50:50:3) with increasing amount of methanol was used as an eluent. Pure fractions were pooled and extracted with 1M phosphate buffer (pH 8.0) and water. Organic phase

was dried with anhydrous Na_2SO_4 and then treated with hexane to effect precipitation. Pure compounds appeared as red powders dried under vacuum.

Complete structural determination of the compounds were accomplished by NMR. Proton and carbon assignments were based on a conventional NOE difference, pHSQC and HMBC measurements. Connectivities in particular relied heavily on HMBC experiment.

As deduced from the data given in Table 4, the structures revealed were L–rhamnosyl–ε–rhodomycinone (**1**) and L–daunosaminyl–ε–rhodomycinone (**2**) shown in Figure 4.

Although these structures were not novel, the generation of the hybrid products by the genes involved in glycosylation portion of aclacinomycin biosynthesis well demonstrates that the genes of pSgs4 are functional and ready to use in drug discovery for finding novel molecules.

Deposited microorganisms

The following microorganisms were deposited according to the Budapest Treaty at Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ), Mascheroder Weg 1b, D–38124 Braunschweig, Germany.

Microorganism	**Accession number**	**Date of deposit**
S. lividans TK24/pSgs4	DSM 12998	12 August 1999
E. coli XL1BlueMRF'/pSgc5	DSM 12999	12 August 1999

Table 4. ^{1}H and ^{13}C chemical shifts of **1** ($DMSO_{d6}$) and **2** (trace of TFA in $DMSO_{d6}$) in 400 and 100 MHz respectively.

	1		2	
Site	^{1}H	^{13}C	^{1}H	^{13}C
1	7.74, 1H, dd, 7.5, 0.9	118.9(d)	7.74, 1H, dd, 7.5, 1.0	119.7(d)
2	7.64, 1H, dd, 8.4, 7.5	136.5(d)	7.68, 1H, dd, 8.1, 7.5	137.4(d)
3	7.22, 1H, dd, 8.4, 0.9	124.1(d)	7.24, 1H, dd, 8.1, 1.0	125.0(d)
4	–	161.8(s)	–	162.6(s)
4–OH	12.00, 1H, s	–	exchange broadened	–
4a	–	115.2(s)	–	115.9(s)
5	–	189.9(s)	–	190.6(s)
5a	–	110.4(s)	–	111.4(s)
6	–	156.2(s)	–	157.1(s)
6–OH	13.41, 1H, s	–	exchange broadened	–
6a	–	135.1(s)	–	135.7(s)
7	5.14, 1H, d, 4.5	70.9(d)	5.15, 1H, d, 3.6	71.3(d)
8A	2.31, 1H, d, 15.1	28.9(t)	2.33, 1H, d, 14.6	34.0(t)
8B	2.14, 1H, dd, 15.1, 4.5	–	2.21, 1H, dd, 14.6, 3.8	–
9	–	70.0(s)	–	70.9(s)
10	4.16, 1H, s	51.2(d)	4.23, 1H, s	51.8(d)
10a	–	134.8(s)	–	136.1(s)
11	–	156.0(s)	–	156.8(s)
11–OH	12.77, 1H, s	–	exchange broadened	–
11a	–	110.8(s)	–	111.1(s)
12	–	185.4(s)	–	186.0(s)
12a	–	132.6(s)	–	133.3(s)
13A	1.73, 1H, dq, 13.9, 7.4	31.7(t)	1.83, 1H, dq, 14.1, 7.3	32.0(t)
13B	1.38, 1H, dq, 13.9, 7.4	–	1.47, 1H, dq, 14.1, 7.3	–
14	1.05, 3H, t, 7.4	6.09(q)	1.13,3H, t, 7.3	6.90(q)
15	–	170.4(s)	–	171.1(s)
16	3.63, 3H, s	51.7(q)	3.70, 3H, s	52.3(q)
1′	5.28, 1H, brs	103.7(d)	5.52, 1H, d, 3.1	100.7(d)
2′	3.83, 1H, d, 5.2	70.9(d)	2.18, 2H, m	27.1(t)
3′	3.44, 1H, dd, 9.0, 5.2	70.8(d)	3.40, 1H, dd, 11.8, 5.1	55.5(d)
4′	3.41, 1H, dd, 9.1, 9.0	72.0(d)	3.98, 1H, brs	67.0(d)
5′	3.77. 1H, dq, 9.1, 6.2	68.9(d)	4.21, 1H, q, 6.3	65.3(d)
6′	1.29, 3H, d, 6.2	16.9(q)	1.32, 3H, t, 6.3	16.7(q)

References

Bibb, M. J., Janssen, G. R., and Ward, J. M. 1985. Cloning and analysis of the promoter region of the erythromycin resistance gene (*erm*E) of *Streptomyces erythraeus*. Gene **38:** 215–226.

El Khamed, H.S., Swartz, D.L., and Cermak, R.C. 1977. Synthesis of ε–rhodomycinone glycosides. J Med Chem **20:** 957–960.

Essery, J.M., and Doyle, T.W. 1980. The synthesis of daunosaminyl–ε–rhodomycinone, daunosaminyl–10–epi–ε–rhodomycinone, daunosaminyl–ε–pyrromycinone and 10–descarbomethoxy–ε–pyrromycin. Can J Chem **58:** 1869–1874.

Hutchinson, C.R., and Fujii, I. 1995. Polyketide synthase gene manipulation: A structure–function approach in engineering novel antibiotics. Annu Rev Microbiol **49:** 201–238.

Kunnari, T., Tuikkanen, J., Hautala, A., Hakala, J., Ylihonko, K., and Mäntsälä, P. 1997. Isolation and characterization of 8–demethoxy steffimycins and generation of 2,8–demethoxy steffimycins in *Streptomyces steffisburgensis* by the nogalamycin biosynthesis genes. J Antibiot **50:** 496–501.

Madduri, K., Kennedy, J., Rivola,G., Inventi–Solari, A., Filippini, S., Zanuso, G., Colombo, A.L., Gewain, K.M., Occi, J.L., MacNeil, D.J., and Hutchinson, C.R. 1998. Production of the antitumor drug epirubicin (4'–epidoxorubicin) and its precursor by a genetically engineered strain of *Streptomyces peucetius*. Nature Biotech **16:** 69–74.

Niemi, J., Ylihonko, K., Hakala, J., Kopio, A., Pärssinen, R., and Mäntsälä, P. 1994. Hybrid anthracycline antibiotics: production of new anthracyclines by cloned genes from *Streptomyces purpurascens* in *Streptomyces galilaeus*. Microbiol **140:** 1351–1358.

Oki, T., Matsuzawa, Y., Yoshimoto, A., Numata, K., Kitamura, I., Hori, S., Takamatsu, A., Umezawa, H., Ishizuka, M., Naganawa, H., Suda, H., Hamada, M, and Takeuchi, T. 1975. New antitumor antibiotics, Aclacinomycins A and B. J Antibiot **28:** 830–834.

Strohl, W. R., Dickens, M. L., Rajgarhia, V. B., Woo, A. J., and Priestley, N. D. 1997. Anthracyclines in Biotechnology of Antibiotics, ed. Strohl, W. R. Marcel Dekker Inc., New York. pp. 577–657.

Strohl, W.R., Bartel, P.L. Li, Y., Connors, N.C., and Woodman, R.H. 1991. Expression of polyketide biosynthesis and regulatory genes in heterologous streptomycetes. J Ind Microbiol **7:** 3: 163–174.

Tsukamoto, N., Fujii, I., Ebizuka, Y., and Sankawa, U. 1994. Nucleotide sequence of the *akn*A region of the aklavinone biosynthetic gene cluster of *Streptomyces galilaeus*. J Bacteriol **176:** 2473–2475.

Ward, J. M., Janssen, G. R., Kieser, T., Bibb, M. J., Buttner, M. J., and Bibb, M. J. 1986. Construction and characterization of a series of multicopy promoter–probe plasmid vectors for *Streptomyces* using the aminoglycoside phosphotransferase from Tn5 as indicator. Mol Gen Genet **203:** 468–478.
Wright, F., and Bibb, M. J. 1992. Codon usage in the G+C–rich *Streptomyces* genome. Gene **113:** 55–65.

Ylihonko K., Hakala J., Kunnari T., and Mäntsälä P. 1996a. Production of hybrid anthracycline antibiotics by heterologous expression of *Streptomyces nogalater* nogalamycin biosynthesis genes. Microbiol **142:** 1965–1972.

Ylihonko, K., Tuikkanen, J., Jussila, S., Cong, L., and Mäntsälä, P. 1996b. A gene cluster involved in nogalamycin biosynthesis from *Streptomyces nogalater*: sequence analysis and complementation of early–block mutations in the anthracycline pathway. Mol Gen Genet **251:** 113–120.

Ylihonko, K., Hakala, J., Niemi, J., Lundell, J., and Mäntsälä, P. 1994. Isolation and characterization of aclacinomycin A–non–producing *Streptomyces galilaeus* (ATCC 31615) mutants. Microbiol **140:** 1359–1365.

SEQUENCE LISTING

<110> Galilaeus Oy

<120> The gene cluster involved in aclacinomycin
biosynthesis, and its use for genetic engineering

<130> 31344

<140>
<141>

<160> 16

<170> PatentIn Ver. 2.2

<210> 1
<211> 662
<212> PRT
<213> *Streptomyces galilaeus*

<400> 1
Met Thr Glu Asp Arg Val Thr Thr Leu Gly Gly Glu Gln Ile Ala Leu
1 5 10 15

Leu Ala Pro Leu Leu Asp Gly Ser Arg Asp Leu Pro Gly Ile Val Ala
20 25 30

Asp Ala Ala Pro Arg Leu Pro Ala Gly Leu Ala Glu Arg Leu Val Thr
35 40 45

Arg Leu Leu Asp Ala Gly Leu Leu Cys Ala Tyr Pro Gln Asp Gly Ala
50 55 60

Asp Arg Pro Glu Arg Ala Tyr Arg Ser Leu Thr Gly Leu Gln Ala Arg
65 70 75 80

Ser Ala Asp Ala Arg Asp Ala Val Leu Ala Ala Val Asp Leu Thr Gly
85 90 95

Asp Ala Glu Ser Pro Leu Pro Glu Ala Val Ser Ala Ala Gly Leu Arg
100 105 110

Ala Ala Ala Pro Gly Glu His Ala Ala Leu Thr Leu Val Leu Cys His
115 120 125

Asp Tyr Leu Asp Pro Arg Leu Ser Ala Leu Asp Ala Glu His Arg Ala
130 135 140

Thr Gly Arg Gly Trp Leu Pro Val Arg Ala Asn Gly Thr His Leu Trp
145 150 155 160

Ile Gly Pro Phe Phe Ser Ala Gly Asp Gly Pro Cys Trp Ser Cys Leu
165 170 175

Ala Asp Arg Leu Arg Leu Arg Arg Arg Gly Glu Ala Tyr Val Gln His
180 185 190

Arg Leu Gly His Ser Gly Pro Ala Val His Arg Arg Ala Tyr Leu Pro
195 200 205

Ala Gly Arg Ala Ala Ala Leu Gln Leu Ala Leu Leu Glu Ala Gly Lys
210 215 220

Trp Leu Ser Gly His Arg Asp Thr Val Gln Asp Ser Leu Trp Arg Leu
225 230 235 240

Asp Thr Arg Thr Leu Glu Ser Ser Arg His Pro Val Arg Arg Arg Pro
245 250 255

Gln Cys Ser Arg Cys Gly Asp Pro Leu Leu Val Arg Asp Arg Val Ser
260 265 270

Ala Pro Val Val Leu Ser Ser Arg Pro Val Arg Asp Glu Ser Gly Gly
275 280 285

Gly His Arg Thr Phe Gly Pro Gln Glu Met Leu Asp Arg Tyr Gly His
290 295 300

Leu Val Asp Pro Val Thr Gly Val Val Gly Glu Ile Arg Arg Asp Pro
305 310 315 320

Arg Gly Pro Glu Phe Leu Asn Cys Phe Thr Arg Ser Arg Cys Arg Leu
325 330 335

Gly Pro Arg Ala Ala Pro Pro Ala Leu His Ser Pro Leu Arg Ser Pro
340 345 350

Gly Ser Gly Lys Gly Val Thr Glu Leu His Ala Arg Val Ser Ala Leu
355 360 365

Ala Glu Ala Leu Glu Arg Cys Ser Gly Tyr Phe Gln Gly Asp Glu Pro
370 375 380

Arg Arg Arg Gly Ser Tyr Arg Glu Leu Ala Gly Leu Ala Val His Pro
385 390 395 400

Asp Ser Val Gln Leu Phe Asp Arg Arg Gln Phe Glu Asp Arg Arg Ala
405 410 415

Trp Asn Arg Ala His Gly Pro Phe His Gln Val Thr Glu Pro Phe Asp
420 425 430

Glu Asp Ala Pro Ile Asp Trp Thr Pro Val Trp Ser Leu Thr Glu Arg
435 440 445

Arg Gln Arg Leu Ala Pro Thr Ser Leu Leu Tyr Tyr Asn Ala Pro Asp
450 455 460

Ala Asp Thr Gly Phe Cys Arg Ala Thr Ser Asn Gly Ala Ala Ala Gly
465 470 475 480

Thr Ser Leu Glu Asp Ala Val Val His Gly Cys Leu Glu Leu Val Glu
485 490 495

Arg Asp Ala Ile Ala Leu Trp Trp Tyr Asn Arg Thr Arg Gln Pro Gly
500 505 510

Val Thr Leu Asp Ala Arg Asp Pro Trp Ile Thr Arg Leu Arg Ala Val
515 520 525

Leu Arg Asp Leu Gly Arg Thr Val Trp Ala Leu Asp Leu Thr Ser Asp
530 535 540

Leu Gly Ile Pro Val Val Ala Ala Val Ser Val Arg Thr Gly Gly Thr
545 550 555 560

Ala Glu Asp Ile Val Leu Gly Phe Gly Ala His Phe Asp Pro Arg Ile
565 570 575

Ala Leu Arg Arg Ala Leu Thr Glu Leu Ser Gln Met Leu Pro Pro Leu
580 585 590

Ala Gln Glu Thr Ala Gly Asp Ala Ser Ala Tyr Thr Gly Thr Asp Pro
595 600 605

Glu Ala Met Arg Trp Phe Arg His Ala Thr Thr Ala Asn Gln Pro Tyr
610 615 620

Leu Leu Pro Ala Ala Arg Arg Ser Ala Arg Pro Pro Ala Ser Leu Arg
625 630 635 640

Pro Pro Arg Asp Ala Ala Ala Gln Ala Gly Ala Leu Val Ala Leu Leu
645 650 655

Arg Arg His Gly Leu Glu
660

<210> 2
<211> 272
<212> PRT
<213> *Streptomyces galilaeus*

<400> 2
Val Asp Ile Trp Leu Leu Gly Pro Leu Thr Ala Glu Val Arg Gly Arg
1 5 10 15

Ser Ile Val Pro Thr Ala Ala Lys Pro Arg Gln Ile Leu Ala Leu Leu
20 25 30

Ala Ile His Ala Asn Arg Val Leu Pro Val Gly Thr Leu Met Glu Glu
35 40 45

Ile Trp Gly Thr Glu Pro Pro Gln Ser Ala Leu Ala Thr Leu His Thr
50 55 60

Tyr Ile Leu Gln Leu Arg Arg Arg Leu Thr Ala Ala Tyr Gly Asp Glu
65 70 75 80

Gly Gly Val Ser Ala Lys Asp Val Leu Val Thr Gln Tyr Gly Gly Tyr
85 90 95

Cys Trp Gln Ala Pro Thr Asp Ser Val Asp Val Pro Arg Tyr Glu Arg
100 105 110

Leu Val Thr Ala Gly Arg Ile Ala Thr Ala Glu Asp Arg Gln Glu Glu
115 120 125

Ala Ser Ala His Phe Arg Glu Ala Leu Ala Leu Trp Arg Gly Ser Ala
130 135 140

Leu Val Asp Val Arg Ile Gly Pro Val Leu Ser Ile Glu Val Ala Arg
145 150 155 160

Leu Glu Glu Ser Arg Leu Gly Val Leu Glu Arg Cys Leu Glu Ala Asp
165 170 175

Leu Arg Leu Gly Arg His Ala Glu Leu Leu Ala Glu Leu Thr Glu Leu
180 185 190

Thr Gly Arg His Pro Leu His Glu Gly Leu His Ala Gln Cys Met Thr
195 200 205

Ala Leu Tyr Arg Ala Gly Arg Ser Trp Gln Ala Leu Asp Val Tyr Gln
210 215 220

Arg Leu Arg Arg Arg Leu Ala Glu Glu Leu Gly Leu Ser Pro Ser Pro
225 230 235 240

Arg Leu Gln Arg Leu Gln Gln Ala Val Leu Ser Ala Glu Pro Trp Leu
245 250 255

Asp Ala Pro Arg Tyr Gly Gly Asp Pro Val Phe Asp Arg Met Ile Ser
260 265 270

<210> 3
<211> 434
<212> PRT
<213> *Streptomyces galilaeus*

<400> 3
Met Thr Ser Asp Thr Lys Ala Leu Val Leu Glu Gln Val Arg Glu Tyr
1 5 10 15

His Arg Gln Gln Gln Pro Gly Asn Phe Gln Pro Gly Val Thr Pro Ile
20 25 30

Leu Ser Ser Gly Ala Val Leu Asp Glu Glu Asp Arg Val Ala Leu Val
35 40 45

Glu Ala Ala Leu Asp Leu Arg Ile Ala Ala Gly Ala His Ser Arg Arg
50 55 60

Phe Glu Ser Lys Phe Ala Arg His Ile Gly Val Arg Lys Ala His Leu
65 70 75 80

Val Asn Ser Gly Ser Ser Ala Asn Leu Leu Ala Leu Ser Ala Leu Thr
85 90 95

Ser Pro Arg Leu Gly Glu Gln Arg Leu Arg Pro Gly Asp Glu Val Ile
100 105 110

Thr Val Ala Gly Gly Phe Pro Thr Thr Val Asn Pro Ile Leu Gln Asn
115 120 125

Gly Leu Thr Pro Val Phe Val Asp Leu Glu Leu Gly Thr Tyr Asn Thr
130 135 140

Thr Val Glu His Val Arg Ala Ala Ile Ser Asp Arg Thr Arg Ala Ile
145 150 155 160

Met Ile Ala His Thr Leu Gly Asn Pro Tyr Gln Val Ala Glu Ile Gln
165 170 175

Gln Leu Ala Thr Glu His Glu Leu Phe Leu Ile Glu Asp Asn Cys Asp
180 185 190

Ala Val Gly Ser Thr Tyr Gln Gly Arg Met Thr Gly Thr Phe Gly Asp
195 200 205

Leu Ala Thr Val Ser Phe Tyr Pro Ala His His Ile Thr Thr Gly Glu
210 215 220

Gly Gly Cys Val Leu Thr Arg Asn Leu Glu Leu Ala Arg Ile Val Glu
225 230 235 240

Ser Phe Arg Asp Trp Gly Arg Asp Cys Trp Cys Glu Pro Gly Glu Asp
245 250 255

Asn Thr Cys Leu Lys Arg Phe Asp Tyr Gln Leu Gly Asn Leu Pro Lys
260 265 270

Gly Tyr Asp His Lys Tyr Ile Phe Ser His Ile Gly Tyr Asn Leu Lys
275 280 285

Ala Thr Asp Leu Gln Gly Ala Leu Ala Leu Ser Gln Leu Asn Lys Leu
290 295 300

Pro Glu Phe Gly Ala Ala Arg Arg Arg Asn Trp Gln Arg Leu Arg Asp
305 310 315 320

Gly Leu Ala Asp Val Pro Gly Leu Leu Leu Pro Val Ala Thr Pro Gly
325 330 335

Ser Asp Pro Ser Trp Phe Gly Phe Val Ile Thr Val Leu Pro Asp Ala
340 345 350

Thr Tyr Thr Arg Arg Asp Leu Val Ala Phe Leu Glu Glu Arg Arg Ile
355 360 365

Gly Thr Arg Arg Leu Phe Gly Gly Asn Leu Thr Arg His Pro Ala Tyr
370 375 380

Leu Gly Thr Pro His Arg Val Ala Gly Asp Leu Arg Asn Ser Asp Ile
385 390 395 400

Ile Thr Glu Gln Ser Phe Trp Ile Gly Val Tyr Pro Gly Ile Thr Glu
405 410 415

Glu Met Thr Asp Tyr Met Arg Glu Ser Ile Val Glu Phe Val Thr Lys
420 425 430

Asn Gly

<210> 4
<211> 329
<212> PRT
<213> *Streptomyces galilaeus*

<400> 4
Met Pro Lys Asp Thr Pro Arg Pro Val Leu Arg Ile Gly Val Leu Gly
1 5 10 15

Cys Ala Asp Ile Ala Val Arg Arg Ile Leu Pro Ala Ile Val Glu His
20 25 30

Pro Ser Val Arg Leu Val Ala Leu Ala Ser Arg Asp Gly Ala Arg Ala
35 40 45

Glu Arg Leu Ala Ala Arg Phe Gly Cys Ala Ala Val Thr Gly Tyr Lys
50 55 60

Ala Leu Leu Asp Arg Glu Asp Ile Asn Ala Val Tyr Val Pro Leu Pro
65 70 75 80

Pro Gly Met His His Glu Trp Val Thr Glu Ala Leu Thr Ala Gly Lys
85 90 95

His Val Leu Val Glu Lys Pro Leu Ser Thr Thr Tyr Ala Gln Ser Val
100 105 110

Asp Leu Val Ala Met Ala Gly Arg Leu Gly Leu Ala Leu Thr Glu Asn
115 120 125

Phe Met Phe Leu His His Ser Gln His Glu Ala Val Arg Ala Met Thr
130 135 140

Gly Glu Ile Gly Glu Leu Arg Val Phe Thr Ser Ser Phe Gly Val Pro
145 150 155 160

Pro Pro His Pro Ser Ser Phe Arg His Asp Ala Arg Leu Gly Gly Gly
165 170 175

Ala Leu Leu Asp Val Gly Val Tyr Pro Leu Arg Ala Ala Gln Leu His
180 185 190

Leu Ala Gly Glu Leu Asp Val Leu Gly Ala Cys Leu Arg Val Asp Glu
195 200 205

Ala Thr Gly Val Asp Val Ala Gly Ser Ala Leu Leu Ser Thr Ala Thr
210 215 220

Gly Val Thr Ala Gln Leu Asp Phe Gly Phe Gln His Ala Tyr Arg Ser
225 230 235 240

Val Tyr Ala Leu Trp Gly Ser Arg Gly Arg Leu Ser Val Pro Arg Ala
245 250 255

Phe Thr Pro Pro Arg Glu His Arg Pro Val Val Arg Ile Glu Gln Gln
260 265 270

Asp Arg Leu Thr Glu Val Thr Leu Pro Ala Asp His Gln Val Gly Asn
275 280 285

Ala Leu Asp Ala Phe Ala Ser Ala Val His Ser Glu Thr Val Arg Ala
290 295 300

Ser Leu Gly Glu Ala Leu Leu Arg Gln Ala Leu Leu Val Glu Gln Val
305 310 315 320

Arg Lys Ala Ala Arg Val Val Ser Gly
325

<210> 5
<211> 323
<212> PRT
<213> *Streptomyces galilaeus*

<400> 5
Met Arg Val Leu Ile Thr Gly Gly Ala Gly Phe Ile Gly Ser His Tyr
1 5 10 15

Val Arg Ser Leu Leu Ala Gly Thr Leu Pro Gly Pro Arg Pro Ser Arg
20 25 30

Val Thr Val Val Asp Leu Leu Thr Tyr Ala Gly Asp Thr Gly Asn Leu
35 40 45

Pro Leu Ala Asp Pro Arg Leu Asp Phe Arg Arg Leu Asp Ile Arg Asp
50 55 60

Leu Asp Ala Leu Leu Thr Val Val Pro Gly His Asp Ala Val Val His
65 70 75 80

Phe Ala Ala Glu Thr His Val Asp Arg Ser Leu Ser Glu Pro Ala Glu
85 90 95

Phe Val Arg Thr Asn Val Leu Gly Thr Gln Ser Leu Leu Glu Ala Ser
100 105 110

Leu Arg Gly Gly Val Gly Thr Phe Val His Val Ser Thr Asp Glu Val
115 120 125

Tyr Gly Ser Ile Ala Gln Gly Thr Trp Thr Glu Glu Ala Pro Leu Leu
130 135 140

Pro Asn Ser Pro Tyr Ala Ala Ser Lys Ala Gly Ser Asp Leu Val Ala
145 150 155 160

Arg Ser Tyr Trp Arg Thr His Gly Leu Asp Val Arg Thr Thr Arg Cys
165 170 175

Ala Asn Asn Tyr Gly Pro Arg Gln His Pro Glu Lys Leu Ile Pro Leu
180 185 190

Phe Val Thr Glu Leu Leu Ala Gly Arg Pro Val Pro Leu Tyr Gly Asp
195 200 205

Gly Gly Asn Val Arg Glu Trp Leu His Val Asp Asp His Cys Arg Ala
210 215 220

Val His Ala Val Leu Thr Gly Gly Arg Pro Gly Glu Ile Tyr Asn Ile
225 230 235 240

Gly Gly Gly Thr His Leu Thr Asn Arg Glu Met Thr Ala Lys Leu Leu
245 250 255

Ala Leu Cys Gly Thr Asp Trp Ser Arg Val Arg Gln Val Pro Asp Arg
260 265 270

Lys Gly His Asp Leu Arg Tyr Ala Val Asp Asp Thr Lys Ile Arg Glu
275 280 285

Glu Leu Gly Tyr Arg Pro Leu Arg Ser Leu Asp Asp Gly Leu Arg Glu
290 295 300

Val Val Asp Trp Tyr Arg Asp Arg Gln Thr His Arg Pro Glu Pro Ala
305 310 315 320

Glu Arg Val

<210> 6
<211> 443
<212> PRT
<213> *Streptomyces galilaeus*

<400> 6
Met Arg Val Leu Leu Thr Ser Phe Ala Leu Asp Ala His Phe Asn Gly
1 5 10 15

Ser Val Pro Leu Ala Trp Ala Leu Arg Ala Ala Gly His Glu Val Arg
20 25 30

Val Ala Ser Gln Pro Ala Leu Thr Ala Ser Ile Thr Ala Ala Gly Leu
35 40 45

Thr Ala Val Pro Val Gly Ala Asp Pro Arg Leu Asp Glu Met Val Lys
50 55 60

Gly Val Gly Asp Ala Val Leu Ser His His Ala Asp Gln Ser Leu Asp
65 70 75 80

Ala Asp Thr Pro Gly Gln Leu Thr Pro Ala Phe Leu Gln Gly Trp Asp
85 90 95

Thr Met Met Thr Ala Thr Phe Tyr Thr Leu Ile Asn Asp Asp Pro Met
100 105 110

Val Asp Asp Leu Val Ala Phe Ala Arg Gly Trp Glu Pro Asp Leu Ile
115 120 125

Leu Trp Glu Pro Phe Thr Phe Ala Gly Ala Val Ala Ala Lys Val Thr
130 135 140

Gly Ala Ala His Ala Arg Leu Leu Ser Phe Pro Asp Leu Phe Met Ser
145 150 155 160

Met Arg Arg Ala Tyr Leu Ala Gln Leu Gly Ala Ala Pro Ala Gly Pro
165 170 175

Ala Gly Gly Asn Gly Thr Thr His Pro Asp Asp Ser Leu Gly Gln Trp
180 185 190

Leu Glu Trp Thr Leu Gly Arg Tyr Gly Val Pro Phe Asp Glu Glu Ala
195 200 205

Val Thr Gly Gln Trp Ser Val Asp Gln Val Pro Arg Ser Phe Arg Pro
210 215 220

Pro Ser Asp Arg Pro Val Val Gly Met Arg Tyr Val Pro Tyr Asn Gly
225 230 235 240

Pro Gly Pro Ala Val Val Pro Asp Trp Leu Arg Val Pro Pro Thr Arg
245 250 255

Pro Arg Val Cys Val Thr Leu Gly Met Thr Ala Arg Thr Ser Glu Phe
260 265 270

Pro Asn Ala Val Pro Val Asp Leu Val Leu Lys Ala Val Glu Gly Leu
275 280 285

Asp Ile Glu Val Val Ala Thr Leu Asp Ala Glu Glu Arg Ala Leu Leu
290 295 300

Thr His Val Pro Asp Asn Val Arg Leu Val Asp His Val Pro Leu His
305 310 315 320

Ala Leu Leu Pro Thr Cys Ala Ala Ile Val His His Gly Gly Ala Gly
325 330 335

Thr Trp Ser Thr Ala Leu Val Glu Gly Val Pro Gln Ile Ala Met Gly
340 345 350

Trp Ile Trp Asp Ala Ile Asp Arg Ala Gln Arg Gln Gln Ala Leu Gly
355 360 365

Ala Gly Leu His Leu Pro Ser His Glu Val Thr Val Glu Gly Leu Arg
370 375 380

Gly Arg Leu Val Arg Leu Leu Asp Glu Pro Ser Phe Thr Ala Ala Ala
385 390 395 400

Ala Arg Leu Arg Ala Glu Ala Glu Ser Glu Pro Thr Pro Ala Gln Val
405 410 415

Val Pro Val Leu Glu Arg Leu Thr Ala Gln His Arg Ala Arg Glu Pro
420 425 430

Arg Arg Pro Gly Gly Thr Ser Pro Cys Val Ser
435 440

<210> 7
<211> 443
<212> PRT
<213> *Streptomyces galilaeus*

<400> 7
Val Gln Thr Gln Asn Ala Pro Glu Thr Ala Glu Asn Gln Gln Thr Asp
1 5 10 15

Ser Glu Leu Gly Arg His Leu Leu Thr Ala Arg Gly Phe His Trp Ile
20 25 30

Tyr Gly Thr Ser Gly Asp Pro Tyr Ala Leu Thr Leu Arg Ala Glu Ser
35 40 45

Asp Asp Pro Ala Leu Leu Thr Arg Arg Ile Arg Glu Ala Gly Thr Pro
50 55 60

Leu Trp Gln Ser Thr Thr Gly Ala Trp Val Thr Gly Arg His Gly Val
65 70 75 80

Ala Ala Glu Ala Leu Ala Asp Pro Arg Leu Ala Leu Arg His Ala Asp
85 90 95

Leu Pro Gly Pro Gln Arg His Val Phe Ser Asp Ala Trp Ser Asn Pro
100 105 110

Gln Leu Cys His Ile Ile Pro Leu Asp Arg Ala Phe Leu His Ala Ser
115 120 125

Asp Ala Asp His Thr Arg Trp Ala Arg Ser Ala Ser Ala Val Leu Gly
130 135 140

Ser Ala Gly Gly Ala Pro Ala Glu Gly Val Arg Glu His Ala Gly Arg
145 150 155 160

Val His Arg Glu Ala Ala Asp Arg Thr Gly Asp Ser Phe Asp Leu Met
165 170 175

Ala Asp Tyr Ser Arg Pro Val Ala Thr Glu Ala Ala Ala Glu Leu Leu
180 185 190

Gly Val Pro Ala Ala Gln Arg Glu Arg Phe Ala Ala Thr Cys Leu Ala
195 200 205

Leu Gly Val Ala Leu Asp Ala Ala Leu Cys Pro Gln Pro Leu Ala Val
210 215 220

Thr Arg Arg Leu Thr Glu Ala Val Glu Asp Val Arg Ala Leu Val Gly
225 230 235 240

Asp Leu Val Glu Ala Arg Arg Thr Gln Pro Gly Asp Asp Leu Leu Ser
245 250 255

Ala Val Leu His Ala Gly Ser Ser Ala Ala Ser Ala Gly Gln Asp Ala
260 265 270

Leu Ala Val Gly Val Leu Thr Ala Val Val Gly Val Glu Val Thr Ala
275 280 285

Gly Leu Ile Asn Asn Thr Leu Glu Ser Leu Leu Thr Arg Pro Val Gln
290 295 300

Trp Ala Arg Leu Gly Glu Asn Pro Glu Leu Ala Ala Gly Ala Val Glu
305 310 315 320

Glu Ala Leu Arg Phe Ala Pro Pro Val Arg Leu Glu Ser Arg Ile Ala
325 330 335

Ala Glu Asp Leu Thr Leu Gly Gly Gln Asp Leu Pro Ala Gly Ala Gln
340 345 350

Val Val Val His Val Gly Ala Ala Asn Arg Asp Pro Glu Ala Phe Leu
355 360 365

Ala Pro Asp His Phe Asp Leu Asp Arg Pro Ala Gly Gln Gly Gln Leu
370 375 380

Ser Leu Ser Gly Pro His Thr Ala Leu Phe Gly Ala Phe Ala Arg Leu
385 390 395 400

Gln Ala Glu Thr Ala Val Arg Thr Leu Arg Glu Arg Arg Pro Val Leu
405 410 415

Ala Pro Ala Gly Ala Val Leu Arg Arg Met Arg Ser Pro Val Leu Gly
420 425 430

Ala Val Leu Arg Phe Pro Leu Thr Thr Ser Ala
435 440

<210> 8
<211> 267
<212> PRT
<213> *Streptomyces galilaeus*

<400> 8
Val Asn Arg Ala Ala Arg Pro Thr Val Arg Gly Met Ser Ala Ile Ala
1 5 10 15

Glu Pro Thr Ala Pro Arg Gly Val Ile Val Thr Gly Gly Gly Thr Gly
20 25 30

Ile Gly Arg Ala Thr Ala His Ala Phe Ala Asp Arg Gly Asp Arg Val
35 40 45

Leu Val Val Gly Arg Thr Ala Ala Thr Leu Ala Gly Thr Ala Glu Gly
50 55 60

His Pro Gly Ile Ser Val Leu Thr Ala Asp Leu Thr Asp Pro Asp Gly
65 70 75 80

Pro Arg Ala Ile Thr Asp Ala Ala Leu Asp Ala Leu Gly Arg Ile Asp
85 90 95

Val Leu Val Asn Asn Ala Ala Thr Gly Gly Phe Ala Gly Leu Ala Glu
100 105 110

Thr Glu Pro Glu Ala Ala Arg Glu Gln Phe Asp Ser Asn Leu Leu Ala
115 120 125

Pro Leu Leu Leu Thr Arg Gln Thr Leu Asp Ala Leu Ser Ala Asp Gly
130 135 140

Gly Gly Thr Val Leu Asn Ile Gly Ser Ala Gly Ala Leu Gly Arg Arg
145 150 155 160

Ala Trp Pro Gln Asn Gly Val Tyr Gly Ala Ala Lys Ala Gly Leu Asp
165 170 175

Phe Leu Thr Arg Thr Trp Ala Val Glu Leu Ala Pro Arg Gly Ile Arg
180 185 190

Val Leu Gly Leu Ala Pro Gly Val Ile Asp Thr Gly Ile Gly Glu Arg
195 200 205

Ser Gly Met Ser Arg Glu Ala Tyr Ala Gly Phe Leu Gly Gln Ile Ala
210 215 220

Ala Arg Val Pro Ala Gly Arg Val Gly Arg Pro Glu Asp Ile Ala Trp
225 230 235 240

Trp Ala Val Gln Leu Ala Asp Pro Arg Ala Ala Tyr Ala Thr Gly Ala
245 250 255

Val Leu Ala Val Asp Gly Gly Leu Ser Leu Thr
260 265

<210> 9
<211> 144
<212> PRT
<213> *Streptomyces galilaeus*

<400> 9
Met Thr Ala Gln Ala Pro Thr Ala Pro Ala Asp Val Tyr Ala Glu Val
1 5 10 15

Gln His Phe Tyr Ala Arg Gln Met Arg Tyr Leu Asp Ser Gly Glu Ala
20 25 30

Glu Thr Trp Ala Gly Thr Phe Thr Glu Asp Gly Ser Phe Ala Pro Pro
35 40 45

Ser Leu Pro Glu Pro Val Arg Gly Arg Pro Leu Leu Ala Glu Gly Ala
50 55 60

Arg Asn Ala Ala Ala Gly Leu Ala Ala Ala Arg Glu Thr His Arg His
65 70 75 80

Trp Val Gly Met Leu Thr Val Thr Pro Ala Asp Asp Gly Ser Leu Thr
85 90 95

Ala Glu Ser Leu Val Ser Ile Val Ala Val Ala Gln Gly Gly Pro Ala
100 105 110

Arg Leu His Leu Val Cys Thr Cys Arg Asp Val Leu Val Arg Glu Gly
115 120 125

Gly Arg Leu Leu Val Arg Glu Arg Val Val Thr Arg Asp Asp Arg Pro
130 135 140

<210> 10
<211> 259
<212> PRT
<213> *Streptomyces galilaeus*

<400> 10
Val Arg Ile Ile Asp Leu Ser Ser Pro Val Asp Ala Ala Gly Phe Glu
1 5 10 15

Pro Asp Pro Val Val His Asp Val Leu Gly Pro Lys Glu Ala Ala Thr
20 25 30

His Met Ser Glu Glu Met Arg Glu His Phe Gly Ile Asp Phe Asp Pro
35 40 45

Ala Glu Leu Pro Glu Gly Glu Phe Leu Ser Leu Asp Arg Leu Gln Leu
50 55 60

Thr Thr His Thr Gly Thr His Val Asp Ala Pro Ser His Tyr Gly Thr
65 70 75 80

Arg Ala Ala Tyr Arg Asp Gly Pro Pro Arg His Ile Asp Glu Met Pro
85 90 95

Leu Asp Trp Phe Phe Arg Pro Ala Val Val Leu Asp Leu Ser Asp Gln
100 105 110

Gly Thr Gly Ala Val Gly Ala Asp Val Leu Arg Arg Glu Met Asp Arg
115 120 125

Ile Gly His Thr Pro Ser Pro Met Asp Ile Val Leu Leu Arg Thr Gly
130 135 140

Ala Asp Ala Trp Ala Gly Thr Pro Lys Tyr Phe Thr Asp Phe Thr Gly
145 150 155 160

Leu Asp Gly Ser Ala Val His Leu Leu Leu Asp Leu Gly Val Arg Val
165 170 175

Ile Gly Thr Asp Ala Phe Ser Leu Asp Ala Pro Phe Gly Asp Ile Ile
180 185 190

Thr Arg Tyr Arg Ala Thr Gly Asp Pro Ser Val Leu Trp Pro Ala His
195 200 205

Val Ile Gly Arg Asp Arg Glu Tyr Cys Gln Val Glu Arg Leu Ala Gly
210 215 220

Leu Asp Arg Leu Pro Ala Ala His Gly Phe Arg Val Ala Cys Phe Pro
225 230 235 240

Val Arg Ile Ala Gly Ala Gly Ala Gly Trp Thr Arg Ala Val Ala Leu
245 250 255

Val Asp Glu

<210> 11
<211> 238
<212> PRT
<213> *Streptomyces galilaeus*

<400> 11
Met Tyr Gly Arg Glu Leu Ala Asp Val Tyr Glu Ala Ile Tyr Arg Ser
1 5 10 15

Arg Gly Lys Asp Trp Gly Gln Glu Ala Ala Asp Val Ser Arg Ile Ile
20 25 30

Thr Glu Arg Arg Pro Gly Ala Gly Ser Leu Leu Asp Val Ala Cys Gly
35 40 45

Thr Gly Ala His Leu Ser Val Phe Ser Thr Leu Phe Glu Val Ala Glu
50 55 60

Gly Leu Glu Ile Ala Glu Pro Met Arg Arg Leu Ala Glu Gln Arg Leu
65 70 75 80

Pro Gly Thr Thr Val His Ala Gly Asp Met Arg Asp Phe Arg Leu Pro
85 90 95

Arg Thr Tyr Asp Ala Val Ser Cys Met Phe Cys Ala Ile Gly Tyr Leu
100 105 110

Glu Thr Leu Asp Asp Met Arg Ala Ala Val Arg Ser Met Ala Ala His
115 120 125

Leu Glu Pro Gly Gly Val Leu Val Val Glu Pro Trp Trp Phe Pro Glu
130 135 140

Asn Phe Ile Glu Gly Tyr Val Ala Gly Asp Leu Ala Arg Glu Glu His
145 150 155 160

Arg Thr Ile Ala Arg Ile Ser His Thr Thr Arg Lys Gly Arg Ala Thr
165 170 175

Arg Met Glu Val Arg Phe Thr Val Gly Asp Ala Ala Gly Ile Gln Gln
180 185 190

Phe Thr Glu Ile Asp Val Leu Thr Leu Phe Thr Arg Asp Glu Tyr Thr
195 200 205

Ala Ala Phe Thr Asp Ala Gly Cys Ser Val Glu Phe Leu Glu Asp Gly
210 215 220

Pro Thr Gly Arg Gly Leu Phe Val Gly Val Arg Glu Gln Arg
225 230 235

<210> 12
<211> 291
<212> PRT
<213> *Streptomyces galilaeus*

<400> 12
Met Lys Gly Ile Ile Leu Ala Gly Gly Ser Gly Thr Arg Leu His Pro
1 5 10 15

Ile Thr Val Ser Val Ser Lys Gln Leu Leu Pro Val Gly Asp Lys Pro
20 25 30

Met Ile Tyr Tyr Pro Leu Ser Val Leu Met Leu Ala Asp Ile Arg Glu
35 40 45

Ile Leu Leu Ile Cys Thr Glu Arg Asp Leu Glu Gln Phe Arg Arg Leu
50 55 60

Leu Gly Asp Gly Ser Gln Leu Gly Leu Arg Ile Asp Tyr Ala Val Gln
65 70 75 80

Asn Arg Pro Ala Gly Leu Ala Asp Ala Phe Val Ile Gly Ala Asp His
85 90 95

Val Gly Asp Asp Asp Val Ala Leu Val Leu Gly Asp Asn Ile Phe His
100 105 110

Gly His His Phe Tyr Asp Leu Leu Gln Ser Asn Val His Asp Val Gln
115 120 125

Gly Cys Val Leu Phe Gly Tyr Pro Val Glu Asp Pro Glu Arg Tyr Gly
130 135 140

Val Gly Glu Thr Asp Ala Ser Gly Gln Leu Val Ser Leu Glu Glu Lys
145 150 155 160

Pro Leu Arg Pro Arg Ser Asp Leu Ala Ile Thr Gly Leu Tyr Leu Tyr
165 170 175

Asp Asn Glu Val Val Asp Ile Ala Lys Asn Leu Arg Pro Ser Pro Arg
180 185 190

Gly Glu Leu Glu Ile Thr Asp Val Asn Arg Asn Tyr Leu Ala Arg Gly
195 200 205

Arg Ala Arg Leu Val Asp Leu Gly Arg Gly Phe Ala Trp Leu Asp Ala
210 215 220

Gly Thr Pro Glu Ser Leu Leu Gln Ala Thr Gln Tyr Val Arg Thr Leu
225 230 235 240

Glu Glu Arg Gln Gly Val Arg Ile Ala Cys Val Glu Glu Val Ala Leu
245 250 255

Arg Met Gly Phe Ile Asp Ala Asp Met Cys His Arg Leu Gly Glu Gln
260 265 270

Met Ser Gln Ser Gly Tyr Gly Arg Tyr Val Met Ala Val Ala Arg Glu
275 280 285

Phe Ser Gly
290

<210> 13
<211> 341
<212> PRT
<213> *Streptomyces galilaeus*

<400> 13
Met Thr Thr Leu Val Trp Asp Tyr Leu Gln Glu Tyr Glu Asn Glu Arg
1 5 10 15

Ala Asp Ile Leu Asp Ala Val Glu Thr Val Phe Ser Ser Gly Arg Leu
20 25 30

Val Leu Gly Asp Ser Val Arg Gly Phe Glu Glu Glu Phe Ala Ala Tyr
35 40 45

His Gly Ala Ala His Cys Val Gly Val Asp Asn Gly Thr Asn Ala Ile
50 55 60

Lys Leu Ala Leu Gln Ala Leu Gly Val Gly Pro Gly Asp Glu Val Val
65 70 75 80

Thr Val Ser Asn Thr Ala Ala Pro Thr Val Val Ala Ile Asp Ser Val
85 90 95

Gly Ala Thr Pro Val Phe Val Asp Val His Pro Asp Ser Tyr Leu Met
100 105 110

Asp Thr Glu Gln Val Glu Ala Ala Leu Thr Pro Arg Thr Arg Cys Leu
115 120 125

Leu Pro Val His Leu Tyr Gly Gln Cys Val Asp Leu Ala Pro Leu Glu
130 135 140

Arg Leu Ala Ala Glu His Asp Leu Phe Leu Val Glu Asp Cys Ala Gln
145 150 155 160

Ala His Gly Ala Arg Arg Ala Gly Arg Leu Ala Gly Thr Thr Gly Asp
165 170 175

Ala Ala Ala Phe Ser Phe Tyr Pro Thr Lys Val Leu Gly Ala Tyr Gly
180 185 190

```
ggcggaccgc gcctgaagac cggtgaggga ccggtaggcc cgttcggggc ggtcggcccc 1800
gtcctgggga taggcgcaca gcagtcccgc gtccagcagc cgggtcacga gccgctcggc 1860
gagtccggcc ggcagccggg gtgccgcgtc ggcgacgatg cccggcagat cgcggctgcc 1920
gtcgagcagc ggggccagca gggcgatctg ctccccgccc agcgtggtca cccggtcctc 1980
ggtcatcagg tagacgggct ctcccggccg cgactcgacc cgcagatgcg gtgcgaaccc 2040
cagccggggg agcccgttgc caggggccgt acggctcatc gagacgcagg gccgccgacc 2100
gtggccgacg gggtgctgga aaggggagcc cagcagatac agatcgcgct gtccgcgtgc 2160
tccccggac cgttgccgaa ctccgtgatg acgagatcgt cggcgaactc ggcggctgcc 2220
gcgtcacctt ccaggagggg cgtgatccgt gacatgtgtg ctccttgtcg ctgtcggccg 2280
gctgctgtga gtagtgctct ggccctggtt cgagctgtgc actgccgtca tccttacggc 2340
ccgtcagggc ggctgccagt gcgtctggca taccgtgcct atgaggttta cacatcttgc 2400
acatacgttc ctcatgtgcc tgttcgggtt cagggcactg gttgattgcc gaagtggcca 2460
gcaagactcc ctgagcatgg agctcttcgc ggtgctccgt cgaccagatg gggagaaga 2520
acgtggacat atggttgctg ggaccgctga cggccgaggt gcggggcagg tcgatcgtgc 2580
ccaccgcggc gaaaccccga cagatcctcg ccctgctcgc catccacgcc aatcgcgtcc 2640
tgcccgtcgg gaccctgatg gaggagatct ggggcaccga gccgccccag agcgccctcg 2700
ccaccctgca cacgtacatc ctccagctgc gccgccggct gaccgctgcc tacggtgacg 2760
agggggcgt gtccgccaag gacgtcctcg tcacccagta cggcggctac tgctggcagg 2820
cgcccacgga ctccgtggac gtaccgcgct acgaacggct cgtcaccgcc ggacggatcg 2880
ccaccgccga ggaccgccag gaggaggcgt cggcccactt ccgtgaggca ctcgcgctct 2940
ggcgggggtc cgcgctggtg gacgtgcgga tcggaccggt cctgagcatc gaggtggcgc 3000
ggctggagga gagcaggctc ggcgtgctgg agcgctgcct ggaggcggac ctgaggctgg 3060
gacgccacgc ggagctgctg gccgaactca ccgaactcac cgggcgccat ccgctgcacg 3120
agggcctgca cgcccagtgc atgacggcgc tgtaccgggc gggccgctcc tggcaggcgc 3180
tggacgtcta ccagaggctg cgccgccggc tggcggagga actcggactc tccccgtcgc 3240
cgcgcctgca gcgtctgcag caggcggtgc tctcggcgga gccctggctg gacgcgccca 3300
ggtacggagg ggacccggtg ttcgaccgga tgatcagctg accgtcccgg cggctcagcc 3360
gttcttggtg acgaactcga cgatggactc ccgcatgtag tcggtcatct cctccgtgat 3420
gcccggatac accccgatcc agaagctctg ctcggtgatg atgtcggagt tgcgcagatc 3480
gcccgccacc cggtgcggcg tgccgaggta cgccggatgg cgggtgaggt tgccgccgaa 3540
cagccgccgt gtgccgatcc gccgttcctc caggaaggcg accaggtcac ggcgggtgta 3600
ggtggcgtcg ggcaggaccg tgatcacgaa cccgaaccag ctgggatcgc tgcccggtgt 3660
cgcgaccggc agcagcagac cggggacgtc ggcgagcccg tcccgcagcc gctgccagtt 3720
gcgccgccgg gccgcgccga actccggcag cttgttcaac tggctgagtg ccagcgcccc 3780
ttgcagatcg gtcgccttca ggttgtaacc gatgtgggag aagatgtact tgtggtcgta 3840
gcccttgggg agattaccca gctggtagtc gaaccgcttg aggcaggtgt tgtcctcgcc 3900
cggttcgcac cagcagtccc ggccccagtc gcggaacgac tcgacgatgc gggccagttc 3960
gaggttgcgg gtcagtacgc agccccccctc gcccgtggtg atgtggtggg caggatagaa 4020
actgaccgtt gccaggtccc cgaaggtccc cgtcatacgt ccctggtagg tcgagcccac 4080
cgcgtcgcag ttgtcctcga tcaggaacag ctcgtgctcg gtggccagtt gctggatctc 4140
cgccacctgg taggggttgc cgagggtgtg cgcgatcatg atggcccggg tccggtccga 4200
gatcgccgcc cgcacgtgct ccaccgttgt gttgtacgta ccgagttcca ggtccacgaa 4260
cacgggcgtc agcccgttct ggaggatcgg gttcaccgtg gtggggaatc cgccggccac 4320
cgtgatcacc tcgtcacccg gacgcagccg ctgctcaccg agccggggag aggtcagcgc 4380
agagagcgcc agcaggttcg ccgaggaccc cgagttcacc aggtgcgcct tgcgtacgcc 4440
gatgtggcgg gcgaatttgc tctcgaagcg ccgggagtgg gctccggccg cgatccgcag 4500
atccagggcg gcctccacca gggcgacccg gtcctcctcg tccaggacgg cgcccgagga 4560
gaggatgggc gtcactccgg gctggaaatt ccctggctgc tgctgccggt gatactcgcg 4620
aacctgctcc aggaccagag ccttggtgtc cgacgtcata gccgtccctc cgtaaggatt 4680
gcctcgccga ccatgctggc ccggatcggt cgaggagccg tggagtcccg gccgaaggcc 4740
ctgctccagt gggcttcaag gggccttcga cctgccgtgg ctagcgtgtc cggcgacgca 4800
tcgaactggt gaggtggcag atgccgaagg acactccacg gcccgtactc cgcatcgggg 4860
ttctgggctg tgccgatatc gcggtgcgcc ggatcctgcc cgcgatcgtg gagcatccgt 4920
cggtccggct ggtcgctctg gcgagccggg acggggcgcg cgccgaacgg ctcgcggccc 4980
gtttcggatg cgcggcggtg accggctaca aggcgctgct ggaccgtgag gacatcaacg 5040
ccgtctacgt tcccctgccg cccggcatgc accacgaatg ggtcaccgaa gcgctgacgg 5100
cgggcaagca cgtgctggtg gagaagccgc tcagcacgac gtacgcgcag agcgtcgacc 5160
tggtggcgat ggccggccgg ctcggcctcg cgctcaccga gaacttcatg ttcctgcacc 5220
actcgcagca cgaggcggtc cgggccatga ccggcgagat cggggaactg cggtcttca 5280
ccagttcctt cggcgtgccg ccgcccccacc cctcgtcctt ccggcacgac gcgcggctcg 5340
gcggcggcgc cctgctggac gtcggtgtct atccgctgcg cgcggcccag ctccacctcg 5400
ccggggaact cgacgtgctg ggcgcctgtc tgcgcgtgga cgaggcgacc ggcgtcgacg 5460
tcgcgggaag cgcgctgctg tccacggcga cgggtgtgac cgcgcagctc gacttcggct 5520
tccagcacgc gtaccggtcc gtgtacgcgc tgtggggcag ccgcggcagg ctgagcgtgc 5580
cgcgggcctt caccccgccc cgtgagcacc gcccggtggt ccgtatcgaa cagcaggacc 5640
gtctcaccga agtgacgctg cccgccgatc accaggtggg caacgcgctc gacgcgttcg 5700
cctcggcggt gcactcggag accgtccgtg cctccctggg ggaggcgctg ctgcgtcagg 5760
```

cgctcctggt cgagcaggta cgcaaagccg cgcgggtcgt cagcggctga gccccccgga 5820
cgctttgcgg gcgcctgacg cccgcacgac gagagnnnnn nnnnnnnnnn nnnnngggct 5880
ctcctcacac tcctcgcggt cgcgccccgc cggggcggct cagacccgct cggccggttc 5940
cggccggtgg gtctgccggt cgcggtacca gtcgacgacc tcgcgcagac cgtcgtccag 6000
ggaccgcagc ggccggtagc cgagttcttc gcggatcttc gtgtcgtcca ccgcgtaccg 6060
caggtcgtgg cccttgcggt cgggcacctg gcggaccctg gaccagtcgg tgccgcacag 6120
cgcgagcagc ttggccgtca tctcacggtt cgtcagatgg gtaccgccgc cgatgttgta 6180
gatctcgccg ggccggccac cggtcagcac cgcgtggacg gcccggcagt ggtcgtccac 6240
gtgcagccat tccctgacgt tgcctccgtc gccgtacagc ggcaccggcc ggccggccag 6300
cagttcggtg acgaacagcg ggatgagctt ctccgggtgc tgacgggggc cgtagttgtt 6360
ggcgcagcgg gtcgtgcgca catcgaggcc gtgggtgcgc cagtaggacc gtgcgaccag 6420
gtcgcttccc gccttcgagg ccgcgtaggg ggagttgggg agcagcggcg cctcctccgt 6480
ccaggtgccc tgcgcgatcg atccgtagac ctcgtccgtg gagacgtgca cgaacgtgcc 6540
gaccccgccg cgcaggctcg cctccagcag cgactgggtg cccagtacgt tggtgcggac 6600
gaactccgcc ggctcgctca gcgaacggtc gacatgggtc tcggccgcga agtgtacgac 6660
cgcgtcgtgg ccggggacca ccgtcagcag cgcgtccagg tcgcggatgt cgagccggcg 6720
gaagtccagc cggggatcgg ccaggggcag attgccggtg tctcccgcgt acgtgagcag 6780
gtcgacgacg gtgacccggc tcgggcgggg accgggcagg gtccccgcca gcagggaacg 6840
gacgtagtgg gatccgatga aaccggcgcc gccggtgatc aggacacgca tggcgacgta 6900
cctccggggc gtcggggttc acgggcgcgg tgctgggcgg tcagccgttc cagtacgggg 6960
acgacctgcg ccggggtcgg ttcggactcg gcctcggccc gcaggcgggc ggccgccgcg 7020
gtgaacgagg gctcgtcgag gagccggacg agcctgccgc gcagtccctc caccgtcacc 7080
tcgtgcgagg ggagatggag gcccgcgccg agtgcctgct gacgctgcgc gcggtcgatc 7140
gcgtcccaga tccagcccat ggcgatctgc ggcacaccct cgaccagcgc cgtcgaccag 7200
gtgccggcgc cgccgtggtg gacgatggcc gcgcaggtcg gcagcagcgc gtgcagcggc 7260
acatggtcca ccagacgcac attgtcgggg acatgggtca gcagtgcccg ttcctcggcg 7320
tcgagggtgg cgaccacctc gatgtccagc ccctcgaccg ccttgaggac gaggtcgacc 7380
gggacggcgt tggggaactc cgaggtccgg gccgtcatgc ccagggtgac gcagacgcgg 7440
gggcgggtcg gcgggacgcg cagccagtcg ggtacgacgg caggccccgg cccgttgtag 7500
gggacgtagc gcatgcccac gaccgggcgg tccgacggcg ggcggaagct gcgcggcacc 7560
tggtccaccg accactgtcc cgtcaccgcc tcctcgtcga aggggacgcc gtagcggccg 7620
agcgtccact ccagccactg gccgagtgag tcgtcggggt gcgtcgtccc gttccccccg 7680
gccgggcccg ccgggggcggc gccgagctgg gcgaggtagg cccgccgcat ggacatgaag 7740
aggtccggaa aggacagcag ccgggcgtgg gccgcgccgg tcaccttcgc ggccaccgcg 7800
ccggcgaagg tgaacggctc ccacaggatc aggtccggtt cccagccgcg ggcgaacgcg 7860
acgaggtcgt cgaccatcgg atcgtcgttg atcagggtgt agaaggtggc ggtcatcatc 7920
gtgtcccagc cctggagaaa ggccggggtg agctgtccgg gggtgtcggc gtcgagcgac 7980
tggtcggcgt ggtgggacag cacagcgtcg cccacgccct tgaccatctc gtccagccgg 8040
gggtccgcgc ccaccggcac ggcggtcagc ccggccgccg tgatgctcgc cgtcagcgcg 8100
ggctggctcg ccacccgcac ctcgtgcccg gcggcccgca acgcccaggc cagcggcacg 8160
gacccgttga agtgagcgtc cagcgcgaag gatgtcagga gaacccgcat ggtcggtcct 8220
ttctctgcgg ctcggctttc ggcggctcgc aactcgggcc gcggcacgcc gggccccggc 8280
gtgccgtcac gcggaggtgg tcagggggaa gcgcagcacc gcgcccagca cgggggaacg 8340
catccgccgc aggaccgcgc cggccggcgc gagcacgggc cggcgctccc gcagcgtgcg 8400
gaccgcggtc tccgcctgga gccgggcgaa cgcccccgaag agcgcggtgt gcgggccgga 8460
gagcgacaac tgcccctgcc cggcgggacg gtcgaggtcg aagtggtccg gggccaggaa 8520
cgcctcgggg tcccggttcg ccgcaccgac atggacgacg acctgggcgc cggccgggag 8580
gtcctggccg cccagggtga ggtcctcggc ggcgatacgg ctctccagcc gcaccggcgg 8640
cgcgaagcgc agtgcctcct ccacggcgcc ggccgcgagt tcagggttct cgccgagccg 8700
cgcccactgc acgggacggg tgagcagcga ctccagggtg ttgttgatga gcccggcggt 8760
gacctcgacg ccgaccacgg cggtcagcac acccacggcg agcgcgtcct gcccggcgga 8820
ggcggcggag gagccggcgt gcagcaccgc gctgagcagg tcgtcacccg gctgcgtccg 8880
ccgcgcctcg acgaggtcgc cgacgagggc ccgtacgtcc tcgacggcct cggtgagccg 8940
gcgggtcacg gcgagcggct gggggcagag cgcggcgtcg agggcgacac cgagggcgag 9000
gcacgtcgcc gcgaagcgtt cccgctgggc ggcggggaca ccgagcagct ccgcggcggc 9060
ctccgtggcc accgggcggg agtagtccgc catgaggtcg aagctgtccc cggtacggtc 9120
ggcggcctcc cggtgcacgc gcccggcgtg ctcgcggaca ccctccgccg gggcgccgcc 9180
cgcactgccc aggacggcgg aggcggaccg cgcccagcgc gtgtggtcgg cgtccgacgc 9240
gtggaggaag gcccggtcca gcgggatgat gtggcacagc tgcggattgc tccaggcgtc 9300
ggagaagacg tgccgctgcg ggcccggcag gtcggcgtgg cgcagggcga gccggggatc 9360
ggccagggcc tcggcggcga cgccgtgccg gccggtgacc caggcgccgg tggtgctctg 9420
ccacagcggc gtcccggctt cccgtatccg gcgggtcagc agcgcggggt cgtcgctctc 9480
ggcgcgcagg gtcagggcgt acgggtcgcc gctggtgccg tagatccagt ggaagccgcg 9540
ggcggtcagg agatgacggc ccagttcgct gtcggtctgc tggttctccg ccgtctccgg 9600
cgcgttctgt gtctgcacgg cagtctcctt cgggacgtgg ttcaggtgag ggacagaccg 9660
ccgtcgaccg cgagcaccgc tccggtcgcg taagcggcgc gcgggtcggc cagctggacg 9720
gcccaccagg cgatgtcctc gggccggccg acccggccgg ccgggacgcg ggcggcgatc 9780

tgcccgagga agcccgcgta ggcctcccgt gacatgcccg agcgttcacc gatgccggtg 9840
tcgatcacgc cgggcgccag gccgaggaca cggatgccgc ggggtgccag ttcgacggcc 9900
caggtccggg tgaggaagtc gagcccggcc ttggcggcgc cgtagacacc gttctgcggc 9960
caggcccggc ggcccagagc tccggccgag ccgatgttca gcacggttcc cccgccgtcg 10020
gccgacaggg cgtccagggt ctgccgggtg agcagcagcg gggccaggag gttgctgtcg 10080
aactgctcgc gggcggcctc cggctcggtc tccgccaggc cggcgaaccc accggtggcc 10140
gcgttgttca cgaggacgtc gatacggccc aacgcgtcca gcgccgcgtc ggtgatggcg 10200
cgcggaccgt ccggatcggt gagatcggcg gtcaggacac tgatgccggg gtggccctcg 10260
gcggtcccgg ccagggtcgc ggcggtgcgg cctacgacga gtacccggtc gccgcggtcg 10320
gcgaaggcgt gggccgtggc ccgcccgatc ccggtcccgc cgccggtgac gatcacgccc 10380
cgcggcgctg tcggttctgc tatggcgctc atgccccgga ccgtaggccg ggccgctcga 10440
ttcacggtcg actcccgctc ggccgcggcc tcagggccgg tcgtcacggg tcaccacgcg 10500
ttcccggacc aggagccggc cgccctcccg gaccaggacg tcccggcagg tgcagaccag 10560
gtgcagccgg gcggggccgc cctgcgcgac ggcgacgatc gagaccaggc tctcggcggt 10620
cagcgagccg tcgtccgcgg gggtgacggt gagcatgccg acccagtggc ggtgcgtctc 10680
gcgggcggcg gccagacccg ccgcggcgtt gcgggcaccc tcggcgagca gcgggcgtcc 10740
gcgtaccggc tccgggagcg agggcggtgc gaaggagccg tcctcggtga aggtgccggc 10800
ccaggtctcg gcctcgcccg agtccaggta ccgcatctgc cgtgcgtaga aatgctgtac 10860
ctcggcgtag acgtccgccg gggccgtcgg tgcctgtgct gtcatgccgt gtcacctttc 10920
gccgttcccg cgtcggggac gggtgcgtgc ccgggtcccc ggtgagtgtc acgacggtcg 10980
cagcgcgact ccaggatcgg tcgagcgggg ggcgggcgga tgcgctggcg gctgttcagt 11040
cagcctccaa cgcttctcaa gccgcgctgg tgagcatggg ctggctcgca ggaacccatc 11100
gccgggaggc gcgtgtgcgc atcatcgacc tgtcctcacc cgtggacgcg gcgggttttg 11160
aacctgatcc cgtcgtgcac gacgttctcg gtccgaagga ggccgccacg cacatgagcg 11220
aggagatgcg tgagcacttc ggcatcgact tcgatccggc ggaactgccg gagggcgaat 11280
tcctctcgct cgaccgtctc cagctgacca cccacaccgg aacccatgtc gacgcgccct 11340
cgcactacgg cacgcgcgcc gcgtaccggg acggtccgcc gcggcacatc gacgagatgc 11400
cgctcgactg gttcttccgg cccgccgtcg tcctcgacct cagcgaccag ggcacgggcg 11460
cggtcggcgc cgacgtgctg cggcgggaga tggaccggat cggccacact ccctcgccca 11520
tggacatcgt cctgctcagg accggtgccg acgcgtgggc gggaaccccg aagtacttca 11580
cggacttcac cgggctggac ggttcggccg tgcacctgct gctggacctg gggtgcggg 11640
tgatcggcac ggacgcgttc agcctggacg cgccgttcgg cgacatcatc acccgctacc 11700
gggccacggg cgacccgtcg gtcctgtggc ccgcccatgt catcgggcgg gaccgggagt 11760
actgccaggt cgagcggctc gccgggctcg accggctgcc cgccgcgcac gggttccggg 11820
tcgcgtgctt cccggtgcgg atcgccggag cgggcgccgg ctggacgagg gcggtggccc 11880
tggtcgacga gtgaggagcg cacggcgggc cgagcggacg acgcgcccac cggggcgccg 11940
acggagggaa cagacatgta cggccgggaa ctcgcggacg tgtacgaggc catctaccgc 12000
agccgcggca aggactgggg acaggaggcg gcggacgtct cgcggatcat caccgaacgg 12060
cgtccgggag ccggctcgct gctcgacgtc gcctgtggca cgggcgccca tctgagcgtg 12120
ttcagcacgc tgttcgaggt cgccgagggc ctggagatcg cggagccgat gcggcggctc 12180
gccgagcagc ggctgcccgg caccaccgtg cacgccggcg acatgcgcga cttccggctc 12240
ccgcgcacct acgacgcggt gagctgcatg ttctgcgcca tcggctatct ggagacgctg 12300
gacgacatgc gggccgccgt ccggtcgatg gccgctcatc tggagcccgg cggcgtgctg 12360
gtcgtcgaac cctggtggtt ccccgagaac ttcatcgagg gctatgtcgc gggtgacctg 12420
gcccgcgagg agcaccggac catcgcccgg atctcgcaca ccacccggaa gggccgggcc 12480
acccgcatgg aggttcgctt caccgtgggg gacgccgccg gcatccagca gttcacggag 12540
atcgacgtgc tgaccctgtt caccagggac gaatacaccg ccgcgttcac cgacgccggc 12600
tgttccgttg aattcctgga ggacggaccc accggccgcg gtcttttcgt cggtgttcgc 12660
gaacagcgct gagccggagg aaggtcagcc cgagaattct cgggcgaccg ccatcacata 12720
gcgaccgtaa ccggactgcg acatctgctc accgagtcgg tggcacatgt ccgcgtcgat 12780
gaatcccatg cgcagcgcca cttcttcgac gcaggcgatg cgcacgccct gccgttcttc 12840
gagcgtgcgc acgtactggg tggcctgaag cagcgattcc ggggttccgg cgtcgagcca 12900
ggcgaaaccg cggccgagat cgaccagacg cgctcggccg cgggccaggt aattccggtt 12960
cacatcggtg atctccagtt cgccgcgcgg cgagggccgg aggttcttcg cgatgtccac 13020
cacttcgttg tcgtacagat agagcccggt gatggcgagg tcggaacgcg gccggagcgg 13080
cttctcctcc agggagacga gttgccccga cgcgtcggtc tcgcccacgc cgtagcgctc 13140
cgggtcctcc acggggtagc cgaacagcac gcagccctgg acgtcgtgga cgttgctctg 13200
gagcaggtcg tagaagtggt gtccgtggaa gatgttgtcc ccgaggacca gtgccacgtc 13260
gtcgtcaccc acgtggtcgg cgccgatgac gaaggcgtcg gccagtccgg ccggccggtt 13320
ctgcaccgcg tagtcgatgc gcaggccgag ctgggagccg tcgcccagga gccggcggaa 13380
ctgttccaga tcgcgctcgg tgcagatgag caggatctcc ctgatgtccg ccagcatcag 13440
caccgagagc ggatagtaga tcatcggctt gtcgccgacc ggaagaagct gttttgatac 13500
ggagacggtt atcggatgga gccgtgttcc cgaaccgccg gcgagaataa tccccttcat 13560
gggtatgccc ctgtcctcga tttcttctca tgctaacgac cggattcgct cggcagcaag 13620
ccggaaaatt tctggtccag ggaaaatcga gcggtaatag agggaatttc ggggttgtgg 13680
accggagcga gaatgtgatg tgctgcgagg ggatggtgac gccctgtggc acggcatcgg 13740
cgtggcatcg gcaaggcatc ggcaagggga ggaaaagaat taatgacgac tctcgtctgg 13800

gactacctac aggaatacga gaacgaacgc gccgacattc tggacgccgt ggagacggtc 13860
ttcagctcgg gtcggctcgt cctcggtgac agcgttcgcg gattcgagga ggagttcgcc 13920
gcgtaccacg gcgcggcgca ctgcgtcggt gtcgacaacg gcaccaacgc gatcaagctg 13980
gccctccagg cgctcggcgt cggccccggt gacgaggtcg tcacggtgtc caacaccgcg 14040
gcccccacgg tcgtggccat cgattcggtg ggcgccaccc cggtgttcgt cgacgtccac 14100
ccggacagct acctcatgga caccgagcag gtggaggccg cactcacgcc ccggacccga 14160
tgcctgctgc cggtccatct ctacgggcag tgcgtcgacc tggctccgct ggagcggctc 14220
gccgcggagc acgacctgtt cctcgtcgag gactgcgcac aggcccacgg tgcccgccgc 14280
gccggccggc tcgccggcac caccggcgac gccgccgcct tctccttcta ccccacgaag 14340
gtgctcggcg cctacggcga cggaggcgcc gtggtgacct cccgggacga cacccaccgc 14400
gcgctgcgcc gactgcgcta ctacggcatg gaggagcggt actacgtcgt cggcaccccg 14460
ggtcacaacg cccggctcga cgaggtccag gccgagatcc tgcggcgcaa gctgcgccgg 14520
ctcgacacct acatcgaggg gcggcgggcc gtcgcccggc gctacgagga cgggctcggc 14580
gacaccggcc tggtgctccc gcacaccgtc cccggcaacg agcacgtgta ctacgtctac 14640
acggtgcgcc acccgcggcg cgacgacatc atcaaggccc tcaaggcgta cgacatcgag 14700
ctgaacatca gctatccctg gcccgtgcac accatgtccg ggttcgccca cctcggctac 14760
ggcaagggct cgctgcccgt caccgaggac ctggccggcc agatct 14806

<210> 15
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: degenerated
oligonucleotide primer

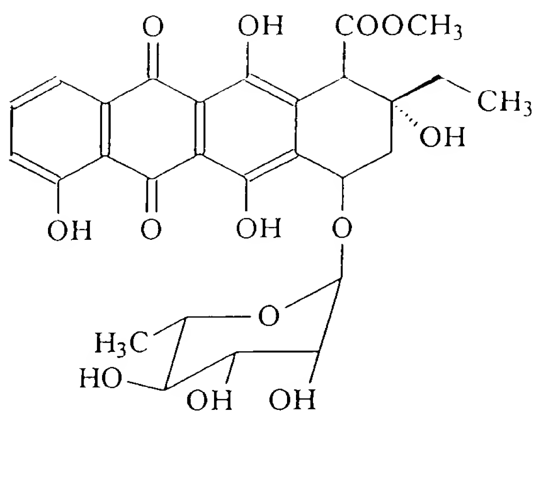
<400> 15
csggsgssgc sggsttcats gg 22

<210> 16
<211> 24
<212> DNA
<213> Artificial Sequence

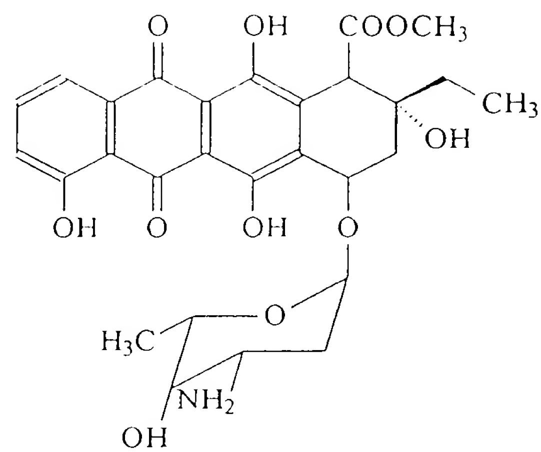
<220>
<223> Description of Artificial Sequence: degenerated
oligonucleotide primer

<400> 16
gggwrctggy rsggsccgta gttg 24

Figure 4



1, L-rhamnosyl-ε-rhodomycinone



2, L-daunosaminyl-ε-rhodomycinone